

# *Evaluation at scale of microbiome-derived metabolites as biomarker of flavan-3-ol intake in epidemiological studies*

Article

Published Version

Creative Commons: Attribution 4.0 (CC-BY)

Open Access

Ottaviani, J. I., Fong, R., Kimball, J., Ensunsa, J. L., Britten, A., Lucarelli, D., Luben, R., Grace, P. B., Mawson, D. H., Tym, A., Wierzbicki, A., Khaw, K.-T., Schroeter, H. and Kuhnle, G. G.C. (2018) Evaluation at scale of microbiome-derived metabolites as biomarker of flavan-3-ol intake in epidemiological studies. *Scientific Reports*, 8 (1). 9859. ISSN 2045-2322 doi: <https://doi.org/10.1038/s41598-018-28333-w> Available at <https://centaur.reading.ac.uk/78125/>

It is advisable to refer to the publisher's version if you intend to cite from the work. See [Guidance on citing](#).

To link to this article DOI: <http://dx.doi.org/10.1038/s41598-018-28333-w>

Publisher: Nature Publishing Group

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the [End User Agreement](#).

[www.reading.ac.uk/centaur](http://www.reading.ac.uk/centaur)

**CentAUR**

Central Archive at the University of Reading

Reading's research outputs online

# SCIENTIFIC REPORTS

OPEN

## Evaluation at scale of microbiome-derived metabolites as biomarker of flavan-3-ol intake in epidemiological studies

Javier I. Ottaviani<sup>1</sup>, Redmond Fong<sup>2</sup>, Jennifer Kimball<sup>2</sup>, Jodi L. Ensunsa<sup>2</sup>, Abigail Britten<sup>3</sup>, Debora Lucarelli<sup>3</sup>, Robert Luben<sup>4</sup>, Philip B. Grace<sup>5</sup>, Deborah H. Mawson<sup>5</sup>, Amy Tym<sup>5</sup>, Antonia Wierzbicki<sup>5</sup>, Kay-Tee Khaw<sup>4</sup>, Hagen Schroeter<sup>1</sup> & Gunter G. C. Kuhnle<sup>4,6</sup>

The accurate assessment of dietary intake is crucial to investigate the effect of diet on health. Currently used methods, relying on self-reporting and food composition data, are known to have limitations and might not be suitable to estimate the intake of many bioactive food components. An alternative are nutritional biomarkers, which can allow an unbiased assessment of intake. They require a careful evaluation of their suitability, including: (a) the availability of a precise, accurate and robust analytical method, (b) their specificity (c) a consistent relationship with actual intake. We have evaluated human metabolites of a microbiome-derived flavan-3-ol catabolite, 5-(3',4'-dihydroxyphenyl)-[gamma]-valerolactone (gVL), as biomarker of flavan-3-ol intake in large epidemiological studies. Flavan-3-ols are widely consumed plant bioactives, which have received considerable interest due to their potential ability to reduce CVD risk. The availability of authentic standards allowed the development of a validated high-throughput method suitable for large-scale studies. In dietary intervention studies, we could show that gVL metabolites are specific for flavan-3-ols present in tea, fruits, wine and cocoa-derived products, with a strong correlation between intake and biomarker (Spearman's  $r = 0.90$ ). This biomarker will allow for the first time to estimate flavan-3-ol intake and further investigation of associations between intake and disease risk.

Flavan-3-ols are a group of phenolic compounds that have been under intensive investigation for their effect on vascular function and the potential to reduce the risk of cardio-vascular diseases<sup>1,2</sup>. Numerous small-scale, short-term studies using surrogate markers of health have been conducted, but it is not clear whether these results provide a reliable estimate of the effect on disease risk reduction and primary disease prevention in the general public<sup>3,4</sup>. This can only be achieved with large, long-term prospective observational epidemiological studies and intervention studies such as COSMOS (NCT02422745). The advantage of the former is that they can provide information about flavan-3-ol intake as part of the habitual diet and that they can make use of data or samples already collected. As of now, epidemiological studies on flavanols, which relied on the validity of the assessment of dietary intake and food content data bases<sup>5</sup>, have shown modest or no relationship between flavanol intake and a decreased risk of cardio-vascular diseases<sup>6</sup>. While dietary intake assessments are often based on a combination of dietary self-reporting and food-composition databases, this approach has well-known limitations<sup>7,8</sup>.

An alternative method for the assessment of dietary intake are nutritional biomarkers. They are objective and measurable characteristics of an individual that are indicative of the intake of a given nutrient or dietary constituent<sup>9</sup>. Nutritional biomarkers can be broadly divided into two categories<sup>10–13</sup>: recovery biomarkers, that allow the accurate estimation of actual intake, such as the use of urinary nitrogen for protein intake<sup>14</sup>, and surrogate or concentration biomarkers that allow the ranking of a study population according to their intake in order to improve dietary assessment or investigate associations with disease risk. Most biomarkers currently used in the context of

<sup>1</sup>Mars, Inc., McLean, VA, USA. <sup>2</sup>Department of Nutrition, UC Davis, CA, USA. <sup>3</sup>MRC Epidemiology Unit, University of Cambridge, Cambridge, UK. <sup>4</sup>Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK. <sup>5</sup>LGC, Newmarket Road, Fordham, UK. <sup>6</sup>Department of Food and Nutritional Sciences, University of Reading, Reading, UK. Correspondence and requests for materials should be addressed to G.G.C.K. (email: [g.g.kuhnle@reading.ac.uk](mailto:g.g.kuhnle@reading.ac.uk))

Time [min]	methanol:acetonitrile (10:90, v/v)	0.1% aqueous formic acid
Initial	5	95
2.00	5	95
3.50	15	85
6.00	20.4	79.6
6.10	95	5
7.00	95	5
7.10	5	95
8.00	5	95
Injection volume	7 µL (nominally – adjusted when required)	
Strong wash	40:30:30 Acetonitrile:Isopropanol:0.1% formic acid (aq)	
Weak wash	95:5 Water:Methanol	

**Table 1.** HPLC-conditions for sample separation (flow rate 0.5 mL/min; column temperature 25 °C).

assessing the intake of bioactives, belong to the latter category<sup>11,12</sup>. These nutritional biomarkers are often based on the concentration of specific metabolites of the compounds of interest in plasma, serum or urine, an approach that we have used previously for isoflavones<sup>15,16</sup>.

In recent years, various human flavan-3-ol metabolites have been identified<sup>17</sup>, and the microbial ring-fission catabolite 5-(3',4'-dihydroxyphenyl)- $\gamma$ -valerolactone (gVL)<sup>18,19</sup>, in particular its phase II metabolites gVL-3'/4'-sulphate and gVL-3'/4'-O-glucuronide (gVLM), have emerged as strong candidates for a nutritional biomarker of flavan-3-ol intake in humans. This is based on their pharmacokinetic characteristics, as their half-life is sufficient to result in steady-state like conditions with regular consumption<sup>17,20</sup>.

Here, we use criteria proposed by IARC<sup>10</sup> and the Institute of Medicine (IOM)<sup>21</sup> to evaluate the suitability of gVLM as a nutritional biomarker of flavan-3-ol intake, in particular that:

- gVLM can be measured accurately and precisely in the relevant human biospecimen, e. g. plasma or urine;
- gVLM are specific for flavan-3-ols;
- gVLM concentration or excretion reflect changes in dietary intake consistently and therefore are prognostic of intake

Furthermore, for use in large-scale epidemiological studies, the analytical method should be robust and suitable for high-throughput analysis.

In this study, we use these criteria to investigate the suitability of gVLM, i.e. the sum of gVL-3'/4'-sulphate and gVL-3'/4'-O-glucuronide, as nutritional biomarkers of flavan-3-ol intake, and investigated whether or not it is suitable for the investigation of flavan-3-ol intake in large cohort studies.

## Methods

**Analytical method.** gVL metabolites were analysed using UPLC-tandem MS. *De novo* synthesised authentic standards<sup>22,23</sup> of the human gVL metabolites 5-(3',4'-dihydroxyphenyl)- $\gamma$ -valerolactone-3'-sulphate, 5-(3',4'-dihydroxyphenyl)- $\gamma$ -valerolactone-3'-O-glucuronide, as well as an isotope labelled internal standard, <sup>13</sup>C<sub>2</sub>D<sub>2</sub>-5-(3',4'-dihydroxyphenyl)- $\gamma$ -valerolactone-3'-sulphate, were obtained from Analyticon (Analyticon Biotechnologies AG, Lichtenfels, Germany). Stock solutions were prepared in ethanol:water (70:30) and stored at −80 °C.

Flavan-3-ol-metabolite-free urine was obtained from human-volunteers on a low flavan-3-ol diet. This study was approved by the University of Reading Research Ethics Committee. All subjects gave their written informed consent to participate and all experiments were performed in accordance with relevant guidelines and regulations.

Spot urine samples (60 µL) and internal standard solutions (2.5 µM <sup>13</sup>C<sub>2</sub>D<sub>2</sub>-5-(3',4'-dihydroxyphenyl)- $\gamma$ -valerolactone-3'-sulphate, 60 µL) were combined, filtered (Impact Protein Precipitation filter plate, Phenomenex, Macclesfield, UK) by centrifugation (500 × g) for 2 minutes at room temperature and stored at −20 °C until analysis. Sample preparation was automated for the analysis of EPIC Norfolk urine samples using a Hamilton Star robot (Hamilton, Bonaduz, Switzerland).

Samples were separated by liquid chromatography (Acquity, Waters, Elstree, UK) using a C18 column (Kinetex C18 200 × 2.1 mm, 1.7 µm, with 0.5 µm Krudcatcher, Phenomenex, Macclesfield, UK) and the conditions described in Table 1.

Compounds were detected by electrospray ionisation tandem mass spectrometry (Applied Biosystems API 4000, Sciex, Warrington, UK) in negative ion mode using the parameters shown in Table 2. The spray voltage was −4500 V and the source temperature was 600 °C.

Samples were quantified using calibration standards prepared in flavan-3-ol-metabolite free urine samples (standard concentrations [µM]: 0.1, 0.25, 10, 20, 30, 40, 50). 5-(3',4'-dihydroxyphenyl)- $\gamma$ -valerolactone-3'-O-glucuronide and 5-(3',4'-dihydroxyphenyl)- $\gamma$ -valerolactone-3'-sulphate were quantified using the peak area ratios of analyte and <sup>13</sup>C<sub>2</sub>D<sub>2</sub>-5-(3',4'-dihydroxyphenyl)- $\gamma$ -valerolactone-3'-sulphate.

Analyte	Precursor ion [m/z]	Product ion [m/z]	Retention time [min]
5-(3',4'-dihydroxyphenyl)- $\gamma$ -valerolactone-3'-sulphate	287	207	3.80
5-(3',4'-dihydroxyphenyl)- $\gamma$ -valerolactone-3'-O-glucuronide	383	207	3.95
<sup>13</sup> C <sub>2</sub> D <sub>2</sub> - $\gamma$ -valerolactone-3'-sulphate	291	211	3.78

**Table 2.** LC-MS parameters for analytes and internal standards. Compounds were detected in negative ion mode with a dwell time of 20 ms.

Values (mg/70 kg BW)	Level 1	Level 2	Level 3	Level 4
Amount of CF	100.0	200.0	400.0	1000.0
Procyanidins	81.7	163.3	326.7	816.7
(-)-Epicatechin	16.0	32.0	64.0	160.0
(+)-Catechin	0.4	0.9	1.8	4.4
(-)-Catechin	1.9	3.8	7.6	18.9
(+)-Epicatechin	0.0	0.0	0.0	0.0
Caffeine	18.1	18.4	19.1	21.1
Theobromine	97.6	97.3	96.9	95.6
Total calories	33.3	33.3	33.3	33.3
Total Carbohydrates	7.8	7.8	7.8	7.8
Sodium	6.7	6.7	6.7	6.7

**Table 3.** Composition of test materials.

Each batch included two replicates of quality control samples with three different concentrations: low QC (0.3  $\mu$ M), medium QC (25  $\mu$ M) and high QC (38  $\mu$ M). Usual acceptance criteria for each batch were that at least one QC at each concentration and four out of the six QCs were within 15% of the theoretical concentration.

The method was validated using flavan-3-ol-metabolite-free urine, spiked with the authentic standards 5-(3',4'-dihydroxyphenyl)- $\gamma$ -valerolactone-3'-O-glucuronide and 5-(3',4'-dihydroxyphenyl)- $\gamma$ -valerolactone-3'-sulphate to assess stability, specificity, matrix effects, precision and accuracy. The matrix effect was assessed by spiking flavan-3-ol-metabolite free urines from different sources with known amounts of analyte and internal standard and comparing their peak-area-ratio ratios. Method performance was assessed using data from quality control samples in 255 batches analysed.

**Human intervention studies.** *Test materials.* Test materials consisted of fruit-flavoured non-dairy drinks that contained either specific flavan-3-ols (specificity study) or varying amounts of flavan-3-ols derived from cocoa (4 levels for the intake amount escalation study). Test materials were freshly prepared on each study day. All test materials were matched regarding their macro- and micro-nutrient content, including theobromine and caffeine, as well as in their orosensory and physicochemical characteristics. Composition of test materials is listed in Table 3. All test materials were supplied by Mars, Incorporated (McLean, VA).

*Study population.* Healthy male adults between 25 and 60 years of age (specificity study) and between 25 and 40 years of age (intake amount escalation study) were recruited by public advertisement in the city of Davis and surrounding areas (California, USA). Exclusion criteria included a body mass index (BMI) higher than 30 kg/m<sup>2</sup>, blood pressure (BP) higher than 140/90 mmHg, allergies to peanut or cocoa, avoidance of caffeinated food products and beverages, a history of CVD, stroke, renal, hepatic, or thyroid disease, gastrointestinal tract disorders, previous gastrointestinal surgery (except appendectomy), the current intake of herbal-, plant- or botanical-containing dietary supplements, persons following vegan/vegetarian diet, and those adhering to an uncommon diet or a weight loss program. To determine eligibility, participants were asked to complete health- and lifestyle questionnaires, have their height, weight, and in-office BP determined, and to provide a blood sample for complete blood count (CBC), liver panel, lipid panel and metabolic panel assessments. Enrolled participants commenced the study protocol between 1–3 weeks after eligibility was determined. While participating in the study, volunteers were asked to maintain their typical daily activities and diet throughout the study. To control for potential dietary flavan-3-ol intake, volunteers were asked to follow a defined low-flavan-3-ol diet on the day prior to and during each study day. All volunteers were instructed on how to follow a low-flavan-3-ol diet, receiving foods containing low or negligible amounts of flavanols including the dinner for the night previous to the study day. Additionally, volunteers were asked to restrain from consuming alcohol, coffee, or any other caffeine-containing beverages for one day prior, and during, the study days. Volunteers were required to fast overnight (12 h water, *ad libitum*) before each study day.

*Specificity study.* In order to investigate the specificity of gVLM, the sum of gVL-3'/4'-sulphate and gVL-3'/4'-O-glucuronide, as nutritional biomarker of flavan-3-ol intake, we performed a dietary intervention study with different possible precursors of gVL. This study was randomized, double-masked and followed an 8x-crossover design, in which volunteers were asked to consume test materials containing specific flavan-3-ols

along with toast with cream cheese. The flavan-3-ols tested were: i) (–)-epicatechin ii) (–)-epigallocatechin, iii) (–)-epicatechin-3-O-gallate, iv) (–)-epigallocatechin-3-O-gallate (all compounds isolated from green tea, purity >99% and food grade), v) a 1:1:1: mixture of theaflavin-3-O-gallate, theaflavin-3'-O-gallate and theaflavin-3,3'-O-digallate, vi) thearubigins (all isolated from black tea, purity >95% and food grade) and vii) procyanidin B-2 ((–)-epicatechin-(4 $\alpha$ 4 $\beta$ )-(–)-epicatechin; isolated from cocoa, purity 91%). One additional visit included the consumption of the test material without any flavan-3-ol added (control day). The amount of (–)-epicatechin consumed per volunteer was 34.8 mg or 120  $\mu$ mol, which is close to the mean intake amount of (–)-epicatechin consumed in the UK<sup>24</sup>. The amount of the rest of the flavan-3-ols was equimolar to the amount of (–)-epicatechin ingested, except for thearubigins, which were given in equal mg amounts to the theaflavin mix.

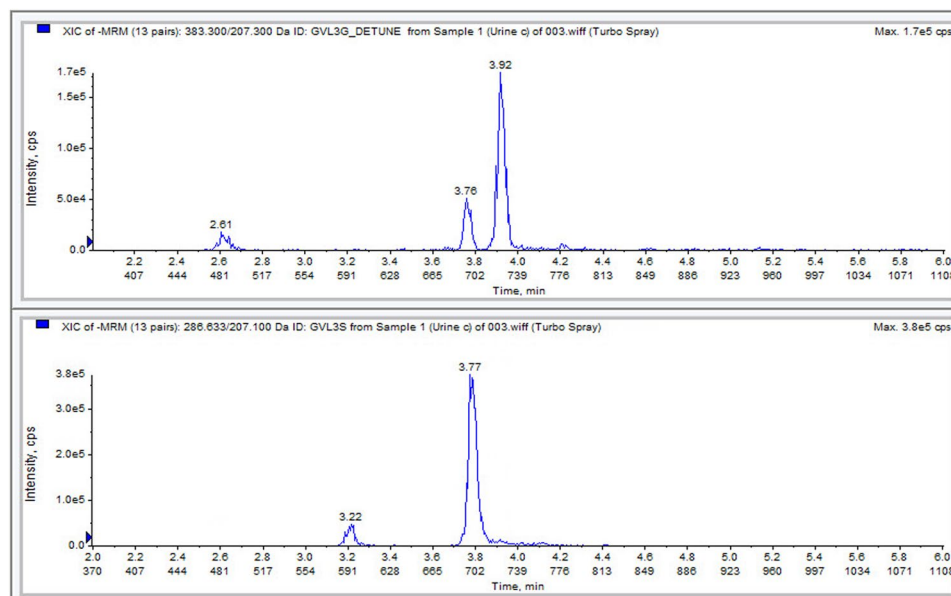
After intake, urine was collected over 24 h in two collection periods (from 0 h to 6 h post intake and 6 h to 24 h post intake), using a fresh container for each collection period with 20 ml of 2 M sodium acetate (pH = 4.5) and 2 mL of 0.5% (w/v) thymol in isopropanol as preservatives. Volunteers returned containers upon completion of sample collection, and urine was stored at –80 °C until analysis. The minimum sample size required to detect a minimum change of 6  $\mu$ mol of gVLM in 24 h-urine (equivalent to a 5% of the amount of flavanols ingested) was determined by pair-t test and imputing a standard deviation of 4  $\mu$ mol. Volunteers were randomized to receive the 8 different test materials using computer-generated lists of random numbers via the randomly permuted block method. The allocation list was generated by a researcher not involved with the recruitment and allocation of participants. We assessed 17 volunteers for eligibility, 12 volunteers were enrolled in the study, and 8 volunteers completed all 8 interventions with the rest of the enrolled volunteers completing <4 interventions. Participants, study nurses, and researchers assessing outcomes as well as researchers involved with the statistical analysis of data were masked to the specific nature of intervention. The initial recruitment started in August 2016, and the study was conducted from August 2016 to January 2017. This study protocol was approved by the Institutional Review Board of UC Davis. All subjects gave their written informed consent to participate and all experiments were performed in accordance with relevant guidelines and regulations. This study was registered as NCT03194620.

**Intake amount escalation study.** In order to investigate the association between flavan-3-ol intake and gVLM excretion, we conducted a dietary intervention study where participants consumed different amounts of flavan-3-ol. This study was randomized, double-masked and followed a 4  $\times$  -crossover design, in which volunteers consumed test materials containing four different amounts of flavan-3-ols from cocoa. These flavan-3-ols include (–)-epicatechin, (+)- and (–)-catechin and their oligomers, the procyanidins, with a degree of polymerization from 2 to 10 in a profile analogous to that naturally found in the seeds of *Theobroma cacao* L. On four separate occasions, each volunteer consumed four different amounts of flavan-3-ols that ranged from 95 mg, including 15 mg of (–)-epicatechin to 1424 mg, including 227 mg of (–)-epicatechin, which are amounts well within the range of intakes determined in the UK<sup>24</sup>. After intake, urine was collected over 24 h in 4 collection periods (from 0 h to 4 h post intake, 4 h to 8 h post intake, 8 h to 12 h post intake and 12 h to 24 h post intake), using a fresh container for each collection period with 20 ml of 2 M sodium acetate (pH = 4.5) and 2 mL of 0.5% (w/v) thymol in isopropanol as preservatives. Volunteers returned the container upon completion of sample collection, and urine was stored at –80 °C until analysis. The minimum sample size required to detect changes of 20  $\mu$ mol of gVLM in 24 h-urine after the intake of each test drink was determined by ANOVA and imputing a standard deviation of 14  $\mu$ mol. Volunteers were randomized to receive the four different test materials using computer-generated lists of random numbers via the randomly permuted block method. The allocation list was generated by a researcher not involved with the recruitment and allocation of participants. We assessed 20 volunteers for eligibility, and 14 volunteers were finally enrolled and completed the study. Participants, study nurses, and researchers assessing outcomes as well as researchers involved with the statistical analysis of data were masked to the intervention. The initial recruitment started in April 2013, and the study was conducted from April to May 2013. This study protocol was approved by the Institutional Review Board of UC Davis. All subjects gave their written informed consent to participate and all experiments were performed in accordance with relevant guidelines and regulations. This study was registered as NCT03201822.

**Sample analysis.** Urine samples from human intervention studies (50  $\mu$ L) were diluted with 100  $\mu$ L of 100  $\mu$ M 3-methyl hippuric acid (internal standard) and analyzed by UPLC-MS within 24 h. Samples were separated on a Waters Acquity UPLC (Waters, Milford, MA, USA) with an autosampler cooled at 5 °C. The injection volume was 37.5  $\mu$ L. Chromatography was carried out at 25 °C using a 100  $\times$  2.1 mm Kinetex C18 1.7  $\mu$ m, reversed phase column with a krudcatcher (Phenomenex). The mobile phase, pumped at a flow rate of 0.5 mL/min, was (A) 5 mM ammonium formate (pH 2.9); and (B) acetonitrile/methanol (90:10, v/v). The gradient started with 5% B from 0–2 min, then ramped to 17.5% from 2–3.5 min, and then to 22.5% from 3.5–6 min. Detection was achieved with a Waters Quattro Micromass spectrometer fitted with an electrospray interface (ESI; Waters). MS analysis was carried out in negative ionization mode by multiple reaction monitoring (MRM). Detection of gVL-3'-sulphate and -3'-O-glucuronide utilized transitions from 287 m/z to 207 m/z and 383 m/z to 207 m/z, respectively, and both with a dwell time of 0.02 s, a cone voltage of 32 V, and a collision energy of 20 V. Data processing was performed using Masslynx software (Waters). MS conditions were set by tuning with the corresponding compound. Source and desolvation temperature were set at 150 °C and 500 °C, respectively, while desolvation and cone gas were set at 900 L/h and 30 L/h, respectively. In both ionization modes, capillary, cone, and lens voltage were set at 4 kV, 30 V, and 1 V, respectively. The identification and quantification of the compounds were based on co-elution with authentic standards. The precision and accuracy of the method was better than 5% and 20%, respectively.

**EPIC Norfolk.** We have investigated the feasibility of using gVLM as a nutritional biomarker in a large observational cohort by applying the method developed in Section 2.1 to 5000 random spot urine samples of the EPIC Norfolk cohort<sup>25</sup>. Participants of this study (n = 25,639), healthy men and women between 40 and 75 years living





**Figure 1.** Typical chromatogram of EPIC Norfolk spot urine sample, showing  $\gamma$ -valerolactone-3'/4'-O-glucuronide (top, 3.9 minutes) and  $\gamma$ -valerolactone-3'-sulphate (bottom, 3.8 minutes) in urine.

in and around Norwich, UK, were recruited between 1993 and 1997. Spot urine samples were collected during the baseline health examination and stored at  $-20^{\circ}\text{C}$  prior to analysis. The study was approved by the Norwich District Health Authority Ethics Committee, and all participants gave signed informed consent.

**Statistical analyses.** Associations between intake and biomarker were assessed with linear regression models using R 3.4.2<sup>26</sup> with the boot package<sup>27</sup>. Adjusted  $R^2$  estimates were obtained by bootstrapping ( $n = 10,000$ ).

**Data availability.** The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

## Results

**Analytical method validation and performance.** The application of a biomarker for nutrient intake assessments requires a robust, validated analytical method, which is ideally suitable for automation and high-throughput analysis. The method developed here can be highly automated and adopted to a robotic sample preparation and analysis platform, which allows for the large-scale analysis of gVLM within an 8 minute run time. Figure 1 shows a typical chromatogram of gVL-3'/4'-sulphate and gVL-3'/4'-O-glucuronide. The precision and accuracy of the method was well within the usual GLP criteria of 15%, and long-term performance monitoring of 255 batches over more than 100 days showed a typical accuracy of better than 5% and a precision better than 10% (Table 4). The estimated limit of detection (signal-to-noise ratio better than 5-to-1) was at least  $0.1\ \mu\text{M}$ . There was a strong matrix effect, which was adequately compensated for by the isotope-labelled internal standard. All compounds were tested for their stability in urine under different conditions, including at high temperature (six hours at  $35^{\circ}\text{C}$ , 24 h hours at room temperature) and following four freeze-thaw cycles ( $-80^{\circ}\text{C}$ ), and the extracted samples were tested for 72 h hours at  $5^{\circ}\text{C}$  and seven days at  $-20^{\circ}\text{C}$ . All compounds were stable under these conditions and showed no indication of deterioration.

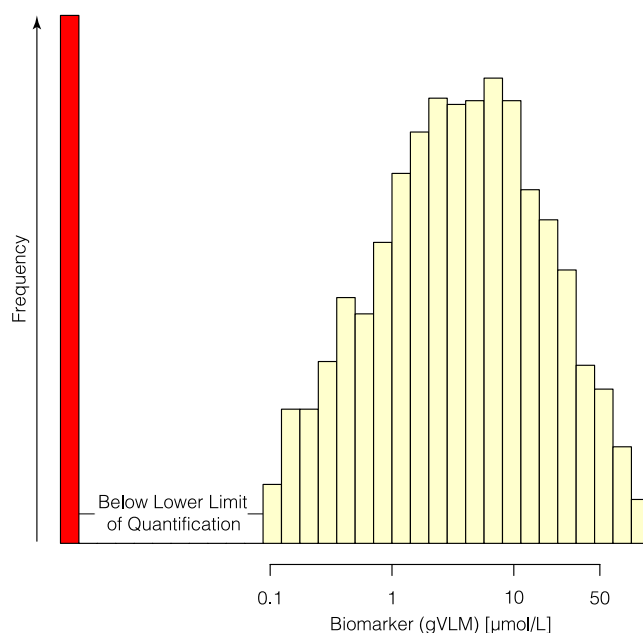
In 5000 random spot urine samples of EPIC Norfolk, median (IQR) concentration of the candidate biomarker (i.e. the sum of gVL-3'/4'-sulphate and gVL-3'/4'-O-glucuronide) was 3 ( $0.8\text{--}10.1$ )  $\mu\text{mol/L}$ , and in 421 (8%) samples the biomarker concentration was below the lower limit of quantification ( $0.1\ \mu\text{mol/L}$ ). The distribution of biomarker concentrations (Fig. 2) was skewed to the right and virtually log-normal.

## Specificity of 5-(3',4'-dihydroxyphenyl)-[gamma]-valerolactone as candidate biomarker of flavan-3-ol intake.

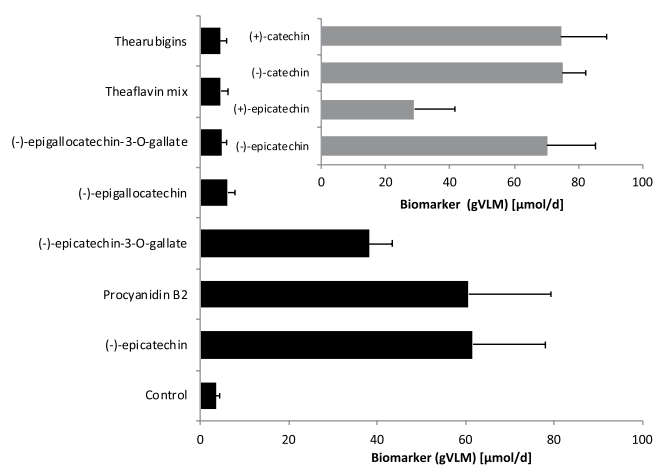
In humans, gVL is formed by the colonic microbiome from different flavan-3-ol precursors<sup>18,19</sup>. However, flavan-3-ols are a complex group of compounds and there was insufficient information about the type of flavan-3-ol that results in the formation of gVL. We have therefore investigated the formation of gVL in humans following the consumption of different monomeric and oligomeric and stereoisomeric flavan-3-ols. This study showed that the oral intake of (–)-epicatechin, (–)-epicatechin-3-O-gallate and procyanidin B-2 resulted in the formation and subsequent presence in urine of both, gVL-3'/4'-sulphate and gVL-3'/4'-O-glucuronide. However, the consumption of theaflavins, thearubigins, (–)-epigallocatechin and (–)-epigallocatechin-3-O-gallate, did not result in the formation of gVL, thus did not contribute to the presence of any of the gVLM in urine. Notably, the intake of equimolar amounts of procyanidin B-2 and (–)-epicatechin-3-O-gallate resulted in similar amounts of gVL-3'/4'-sulphate and –3'/4'-O-glucuronide as (–)-epicatechin (Fig. 3). To further characterize the specificity of gVLM, we have also re-analysed a subset of urine samples collected after the intake of

	Mean (SD) [ $\mu$ M]	%CV	%RE	Mean (SD) [ $\mu$ M]	%CV	%RE	Mean (SD) [ $\mu$ M]	%CV	%RE
	Low QC (0.30 $\mu$ M)			Medium QC (25 $\mu$ M)			High QC (38 $\mu$ M)		
5-(4'-hydroxyphenyl)- $\gamma$ -valerolactone-3'-sulphate	0.29 (0.02)	8.2	-4.6	24.9 (1.1)	4.4	-0.3	37.9 (1.5)	3.9	-0.4
5-(4'-hydroxyphenyl)- $\gamma$ -valerolactone-3'-O-glucuronide	0.30 (0.04)	11.9	-1.1	25.4 (2.2)	8.7	1.9	37.9 (2.9)	7.7	-0.3

**Table 4.** Long-time precision (%CV) and accuracy (%RE, difference of mean calculated concentration and nominal concentration, standardised by nominal concentration) based on 288 analysed batches.



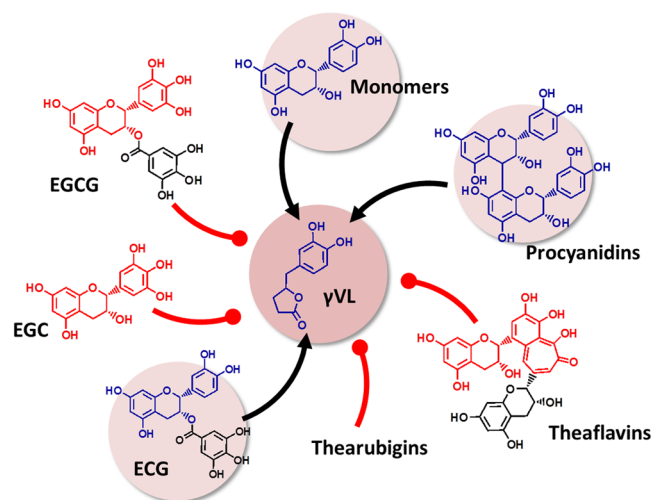
**Figure 2.** Distribution of gVLM (sum of  $\gamma$ -valerolactone-3'/4'-sulfate and  $\gamma$ -valerolactone-3'/4'-O-glucuronide) concentrations in 5000 random samples of EPIC Norfolk. The red bar indicates samples with concentration below the lower limit of quantification (0.1  $\mu$ mol/L).



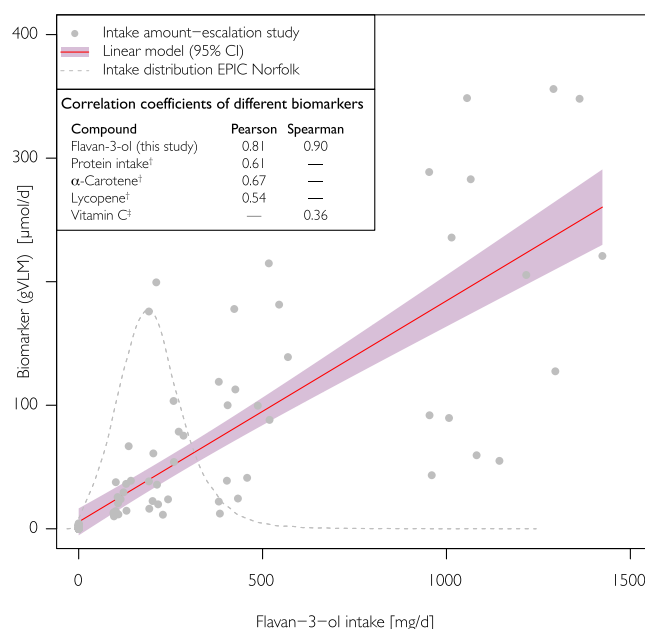
**Figure 3.** Urinary excretion of flavan-3-ol biomarker (gVLM, sum of  $\gamma$ -valerolactone-3'/4'-sulphate and O-glucuronide) following the consumption of different flavan-3-ols (error bars show standard error;  $n = 12-8$ ). \*indicate statistically significant differences ( $p < 0.05$ ; repeated measures ANOVA).

different (epi)catechin stereoisomers in humans<sup>28</sup>, demonstrating that all (epi)catechin stereoisomers also contribute to the presence of gVL-3'/4'-sulphate and -3'/4'-O-glucuronide in urine (Fig. 3, insert). In summary, only flavan-3-ols with an intact (epi)catechin moiety can contribute to the formation of gVL and subsequently urinary gVLM (Fig. 4).





**Figure 4.** Flavan-3-ol precursors of the microbial metabolite 5-(3'/4'-dihydroxyphenyl)- $\gamma$ -valerolactone (gVL). Only compounds with intact (epi)catechin moiety result in the formation of gVL by the intestinal microbiome. ECG, (–)-epicatechin-3-*O*-gallate; ECGG, (–)-epigallocatechin-3-*O*-gallate; EGC, (–)-epigallocatechin.



**Figure 5.** Association between flavan-3-ol intake and gVLM (sum of 5-(3'/4'-dihydroxyphenyl)- $\gamma$ -valerolactone-3'/4'-sulphate and *O*-glucuronide metabolites). Results from the intake-amount escalation study and regression analysis ( $R^2 = 0.66$ ; 95% CI 0.49; 0.80). The intake amount used is comparable with the estimated habitual intake in EPIC Norfolk<sup>6</sup>. For comparison, Pearson and Spearman correlation coefficients for other biomarkers are shown (<sup>†</sup>44; <sup>‡</sup>55).

**Association between flavan-3-ol intake and gVLM.** A strong and consistent association between flavan-3-ol intake and the gVLM is crucial for their use as biomarker of intake. We have therefore investigated this association using daily intakes in humans of up to 1400 mg/d of flavan-3-ols, which covers the entire range of flavan-3-ol intake observed in EPIC Norfolk (0–1200 mg/d<sup>6</sup>). The results of our intake amount-escalation study show a strong and consistent association between flavan-3-ol intake and 24 h urinary excretion of gVLM ( $R^2 = 0.66$ ; 95% CI 0.49; 0.80; Spearman's rank correlation:  $\rho = 0.90$ ; Pearson's  $r = 0.81$ ; Fig. 5). These associations were attenuated, but remained significant, when excluding flavan-3-ol amounts higher than the top decile of intake in EPIC Norfolk (325 mg/d), with  $R^2 = 0.50$  (95% CI 0.34; 0.70), a Spearman's rank correlation coefficient of  $\rho = 0.82$  and a Pearson's  $r = 0.70$ . A similar attenuation was observed when excluding data for a flavan-3-ol intake of 0 mg/d ( $R^2 = 0.49$  (95% CI 0.25; 0.69), Spearman's rank correlation coefficient  $\rho = 0.74$ ; Pearson's  $r = 0.70$ ).

## Discussion

In this study, we have assessed the suitability of the 5-(3',4'-dihydroxyphenyl)- $\gamma$ -valerolactone (gVL) phase II metabolites, gVL-3'/4'-*O*-glucuronide and gVL-3'/4'-sulphate (gVLM), as nutritional biomarker of flavan-3-ol intake, using the criteria suggested by IARC<sup>10</sup> and the Institute of Medicine (IOM)<sup>21</sup>. In the course of this assessment, we (i) developed a robust, precise and accurate method to analyse at scale gVLM in urine samples, (ii) established which flavan-3-ols lead to the formation of gVLM in humans, and (iii) showed a consistent relationship between flavan-3-ol intake and urinary concentrations of gVLM. In the following section, we discuss whether or not gVLM meets each of these criteria.

**Measurement of gVLM as nutritional biomarker.** The application of a nutritional biomarker requires a robust, precise and accurate bioanalytical method to determine the biomarker in the respective biospecimen. Here, we use phase II metabolites of gVL, and their accurate quantification requires the availability of authentic standards as all other methods introduce bias. The most commonly used approach to date, the deconjugation with sulphatases and  $\beta$ -glucuronidase enzymes, followed by the analysis of the aglycon, relies on the complete and consistent hydrolysis of phase II metabolites. However, we and others have shown previously that this cannot be fully and consistently achieved for metabolites of flavan-3-ols<sup>29,30</sup> and that the degree of hydrolysis is affected by reaction conditions<sup>31</sup>. An alternative method not relying on hydrolysis is to identify metabolites by their characteristic tandem-MS fingerprint and to quantify them as aglycone. However, this also introduces large instrument- and method-dependent variability, which renders the results not suitable for biomarker analysis. Authentic standards thus remain the “gold standard” for the identification and quantification of these compounds<sup>32</sup>, and the development of this method was therefore made possible by the availability of authentic standards<sup>22,23</sup>.

Isotope-labelled standards are often used in tandem mass spectrometry for quantification, especially as they can compensate for matrix effects<sup>33</sup>. However, this can be affected by slight differences in retention time between labelled and unlabelled standard<sup>34</sup> due to the isotope effect<sup>35</sup>, resulting in an uncompensated or partially uncompensated matrix effect. For gVLM, the matrix effects were adequately compensated and precision and accuracy for both compounds studied here were well within the accepted range with an error of less than 15%.

The validation of the method developed here was shown to be robust, accurate and precise, and has been validated with reference to generally accepted standards<sup>36</sup>. Furthermore, with a run time of 8-minutes and an easily automatable sample preparation, we show that this method is suitable for robotic automation and therefore for the high-throughput analysis of samples for the application of this method to large-scale studies. The application to a random subset of samples collected from EPIC Norfolk shows that the validation range is sufficient to measure the biomarker in the majority of participants.

**gVLM are specific for flavan-3-ols.** While the relationship between nutritional biomarker and intake is often assessed<sup>11,12</sup>, there is a paucity of data on their specificity. Indeed, it is often tacitly assumed that structurally-related metabolites, such as, for example glucuronidated or sulphated forms, are sufficiently specific for individual compounds<sup>37,38</sup>, but this assumption is rarely proven out. Investigating specificity is especially important when developing a biomarker for a structurally diverse group of compounds, such as flavan-3-ols.

Flavan-3-ols are extensively metabolised by the colonic microbiome<sup>18,39</sup>, but many of the metabolites, such as hippuric acid and phenylacetic acids are also derived from many other precursors<sup>40</sup>, and therefore not suitable to assess their intake. Urinary gVLM in humans has been found only after the intake of flavan-3-ols, and not following the consumption of other flavonoids and polyphenols<sup>17–19</sup>. However, flavan-3-ols encompass a complex group of compounds (Fig. 3), and little was known about the contribution of specific flavan-3-ols to the presence of gVLM in urine. The results obtained here showed that human consumption of (–)-epicatechin, epicatechin-3-*O*-gallate and procyanidin B-2 resulted in the formation of gVL, whereas theaflavins, thearubigins, (–)-epigallocatechin and (–)-epigallocatechin-3-*O*-gallate did not. We also showed that the intake of (–)-epicatechin stereoisomers contributes to the presence of gVL, confirming previous reports<sup>17,18</sup>. The finding that epicatechin-3-*O*-gallate intake gives rise to gVLM in urine implies that epicatechin-3-*O*-gallate is hydrolysed after intake, and that the resulting epicatechin is subsequently converted into gVL by the colonic microbiome. Thus, gallic acid esters of other (–)-epicatechin stereoisomers, like catechin-3-*O*-gallate, are likely to also undergo this transformation. Similarly, other dimeric procyanidins different from procyanidin B-2, like procyanidin B1 [(–)-epicatechin-(4 $\alpha$ )-(+)-catechin], should be expected to contribute to urinary gVLM.

The data available today does not allow for a direct assessment of whether or not procyanidins with a higher degree of polymerisation (DP > 2) contribute to urinary gVLM. However, our data show that only one of the (–)-epicatechin units in procyanidin B2 contributed to the presence of gVLM in urine, and it is likely that this will also be applicable to procyanidins with a higher degree of polymerisation, i.e. that if the terminal unit is a (epi)catechin moiety it gives rise to the formation of gVL. Taken together, our results show that urinary gVLM are specific for the intake of (–)-epicatechin and its stereoisomers, their 3-*O*-gallate esters and procyanidins.

**Relationship between dietary intake and gVLM.** The intended use of a nutritional biomarker is to provide data on dietary intake, and a consistent relationship between dietary intake of a given compound and its biomarker of intake is therefore crucial. This requires an accurate and precise estimate of actual intake, but many studies rely on intake assessments derived from self-reporting methods and food-composition data as reference to validate nutritional biomarkers<sup>41</sup>. This is done despite the known limitations of data derived from self-reporting and food-composition methods<sup>7,8</sup>, especially with regard to a high variability in food content<sup>41,42</sup>. An investigation of the intake-biomarker relationship should therefore either use duplicate diets with validated analytical methods to assess intake, or use defined dietary interventions. In this study, we have used dietary interventions for which detailed compositional data were available (Table 3), and our data show a strong and consistent association between flavan-3-ol intake and urinary gVLM levels. The  $R^2$  value of 0.66 and correlation coefficient

of 0.90 are comparable or better than that of many other well-accepted nutritional biomarkers<sup>43,44</sup>, and only moderately attenuated when excluding upper extremes of intake. Importantly, this strong association was possible even in the context of expected inter-individual differences in the production of gVL<sup>45</sup>. In this manner, gVLM differs from other gut microbiome-derived flavonoid metabolites, like equol, which shows large inter-individual differences, even capable of distinguishing populations as equol producers and non-producers<sup>46</sup>. In this context, and limited by the rather small sample cohort studies, we have previously reported that gVL metabolites show inter-individual differences of 60% in humans<sup>47</sup>. This inter-individual variability of gVL metabolites is more comparable to that reported for other gut microbiome metabolites, like enterolactone and enterodiol<sup>48</sup>, which are well recognized biomarkers of lignan intake<sup>49</sup>.

**Suitability of gVLM as nutritional biomarker of flavan-3-ol intake.** Taken together our results show that gVLM fulfils the criteria set to evaluate candidate nutritional biomarkers, and that gVLM qualifies as a surrogate biomarker<sup>13</sup> of flavan-3-ol intake. It therefore enables the ranking of participants according to their flavan-3-ol intake and thus enables investigations into associations with disease risk biomarkers<sup>10–12</sup> or morbidity/mortality endpoints. While the validation of gVLM has been conducted with 24 h urine samples, surrogate biomarkers can also rely on the use of spot urines. In this context, the plasma half-life of gVLM of approximately 6 h, ( $6.3 \pm 1.7$  h for gVL-3'-sulphate and  $3.1 \pm 0.6$  h for gVL-3'-O-glucuronide)<sup>17</sup>, is sufficient to reach steady-state conditions with regular intake, which is the case for most dietary sources of flavan-3-ols, such as tea and various fruit. Therefore, it can be expected that the urinary excretion of gVLM is sufficiently constant to allow the use of creatinine-adjusted<sup>50</sup> biomarker concentration in spot urine samples. Indeed, data for enterolignans, colonic metabolites with similar half-lives (enterodiol,  $4.4 \pm 1.3$  h; enterolactone  $13 \pm 5.6$  h<sup>51</sup>) to that of gVLM, show very high correlation between spot and 24 h urine concentrations<sup>52</sup>. Frequent high flavan-3-ol intake will therefore result in consistently higher gVLM concentrations in spot urine than low or episodic intake.

The data presented here, in combination with pharmacokinetic data published previously<sup>17</sup>, therefore show that creatinine-adjusted gVLM concentration in spot urine samples can be used as nutritional biomarker of flavan-3-ol intake.

## Conclusion

In this study, we have evaluated the suitability of the microbiome-derived flavan-3-ol metabolite, 5-(3',4'-dihydroxyphenyl)- $\gamma$ -valerolactone (gVL), as nutritional biomarker for the assessment of flavan-3-ol intake in large-scale epidemiological studies. The development of this biomarker has been made possible by the availability of authentic standards to allow the quantification of the Phase II metabolites gVL-3'/4'-sulphate and gVL-3'/4'-O-glucuronide in urine. The method developed is suitable for automation and thus high-throughput analysis, and we have applied it to 5000 random samples of the EPIC Norfolk cohort. The results show that the biomarker is found in the majority of participants with a wide range of concentrations and a high variability, allowing for the ranking of participants according to flavanol-3-ol intake.

We could show that this nutritional biomarker is specific for a number of specific flavan-3-ols present in tea, fruits, wine and cocoa-derived products, which represent the main sources of flavan-3-ols in the diet<sup>53,54</sup>. Overall, the biomarker established here allows for a significantly improved approach in the investigation of the associations between flavan-3-ol intake and health outcomes in large-scale observational studies, thereby providing critical data in the context of developing potential dietary recommendations.

## References

- Schroeter, H. *et al.* Recommending flavanols and procyanidins for cardiovascular health: Current knowledge and future needs. *Molecular Aspects of Medicine* **31**, 546–557 (2010).
- Ottaviani, J. I., Heiss, C., Spencer, J. P. E., Kelm, M. & Schroeter, H. Recommending flavanols and procyanidins for cardiovascular health: Revisited. *Mol. Aspects Med.* <https://doi.org/10.1016/j.mam.2018.02.001> (2018).
- Weintraub, W. S., Lüscher, T. F. & Pocock, S. The perils of surrogate endpoints. *Eur Heart J* **36**, 2212–2218 (2015).
- Bikdeli, B. *et al.* Two Decades of Cardiovascular Trials With Primary Surrogate Endpoints: 1990–2011. *J Am Heart Assoc* **6**, e005285 (2017).
- Yetley, E. A. *et al.* Options for basing Dietary Reference Intakes (DRIs) on chronic disease endpoints: report from a joint US-/Canadian-sponsored working group. *Am J Clin Nutr* **105**, 249S–285S (2017).
- Vogiatzoglou, A. *et al.* Associations between flavan-3-ol intake and CVD risk in the Norfolk cohort of the European Prospective Investigation into Cancer (EPIC-Norfolk). *Free Radic Biol Med* **84**, 1–10 (2015).
- Prentice, R. L. *et al.* Evaluation and comparison of food records, recalls, and frequencies for energy and protein assessment by using recovery biomarkers. *Am J Epidemiol* **174**, 591–603 (2011).
- Subar, A. F. *et al.* Addressing Current Criticism Regarding the Value of Self-Report Dietary Data. *J Nutr* **145**, 2639–2645 (2015).
- Potischman, N. & Freudenheim, J. L. Biomarkers of nutritional exposure and nutritional status: an overview. *J Nutr* **133**(Suppl 3), 873S–874S (2003).
- Kaaks, R., Riboli, E. & Sinha, R. Biochemical markers of dietary intake. *IARC Sci Publ* 103–126 (1997).
- Jenab, M., Slimani, N., Bictash, M., Ferrari, P. & Bingham, S. A. Biomarkers in nutritional epidemiology: applications, needs and new horizons. *Hum. Genet.* **125**, 507–525 (2009).
- Kuhnle, G. G. C. Nutritional biomarkers for objective dietary assessment. *J. Sci. Food Agric.* **92**, 1145–1149 (2012).
- Keogh, R. H., White, I. R. & Bingham, S. A. Using surrogate biomarkers to improve measurement error models in nutritional epidemiology. *Stat Med* **32**, 3838–3861 (2013).
- Bingham, S. A. & Cummings, J. H. Urine nitrogen as an independent validity measure of dietary intake: a study of nitrogen balance in individuals consuming their normal diet. *Am J Clin Nutr* **42**, 1276–1289 (1985).
- Ward, H. *et al.* Breast cancer risk in relation to urinary and serum biomarkers of phytoestrogen exposure in the European Prospective into Cancer-Norfolk cohort study. *Breast Cancer Res* **10**, R32 (2008).
- Ward, H., Chaplais, G., Kuhnle, G. G. & Luben, R. Lack of prospective associations between plasma and urinary phytoestrogens and risk of prostate or colorectal cancer in the European Prospective into Cancer-Norfolk study. *Cancer Epidemiology* **17**, 2891–2894 (2008).
- Ottaviani, J. I. *et al.* The metabolome of [2-<sup>14</sup>C](–)-epicatechin in humans: implications for the assessment of efficacy, safety, and mechanisms of action of polyphenolic bioactives. *Sci Rep* **6**, 1–10 (2016).
- Das, N. P. Studies on flavonoid metabolism. Absorption and metabolism of (+)-catechin in man. *Biochem Pharmacol* **20**, 3435–3445 (1971).

19. Unno, T., Tamemoto, K., Yayabe, F. & Kakuda, T. Urinary excretion of 5-(3',4'-dihydroxyphenyl)-gamma-valerolactone, a ring-fission metabolite of (–)-epicatechin, in rats and its *in vitro* antioxidant activity. *J Agric Food Chem* **51**, 6893–6898 (2003).
20. Borges, G., Ottaviani, J. I., van der Hoof, J. J., Schroeter, H. & Crozier, A. Absorption, metabolism, distribution and excretion of (–)-epicatechin: A review of recent findings. *Mol. Aspects Med.* <https://doi.org/10.1016/j.mam.2017.11.002> (2017).
21. Institute of Medicine (US) Committee on Qualification of Biomarkers and Surrogate Endpoints in Chronic Disease, Mischeel, C. M. & Ball, J. R. *Evaluation of Biomarkers and Surrogate Endpoints in Chronic Disease*. (National Academies Press (US), 2010).
22. Zhang, M., Erik Jagdmann, G. Jr., Van Zandt, M., Beckett, P. & Schroeter, H. Enantioselective synthesis of orthogonally protected (2R,3R)-(–)-epicatechin derivatives, key intermediates in the *de novo* chemical synthesis of (–)-epicatechin glucuronides and sulfates. *Tetrahedron: Asymmetry* **24**, 362–373 (2013).
23. Zhang, M. *et al.* Chemical Synthesis and Characterization of Epicatechin Glucuronides and Sulfates: Bioanalytical Standards for Epicatechin Metabolite Identification. *J Nat Prod* **76**, 157–169 (2013).
24. Vogiatzoglou, A. *et al.* Assessment of the dietary intake of total flavan-3-ols, monomeric flavan-3-ols, proanthocyanidins and theaflavins in the European Union. *Br J Nutr* **111**, 1463–1473 (2014).
25. Day, N. *et al.* EPIC-Norfolk: study design and characteristics of the cohort. European Prospective Investigation of Cancer. *Br J Cancer* **80**(Suppl 1), 95–103 (1999).
26. R Core Team. *R: A Language and Environment for Statistical Computing*. (2017).
27. Canty, A. & Ripley, B. *boot: Bootstrap R (S-Plus) Functions*. (2017).
28. Ottaviani, J. I. *et al.* The stereochemical configuration of flavanols influences the level and metabolism of flavanols in humans and their biological activity *in vivo*. *Free Radic Biol Med* **50**, 237–244 (2011).
29. Saha, S. *et al.* Human O-sulfated metabolites of (–)-epicatechin and methyl-(–)-epicatechin are poor substrates for commercial aryl-sulfatases: Implications for studies concerned with quantifying epicatechin bioavailability. *Pharmacological Research* **65**, 592–602 (2012).
30. Ottaviani, J. I., Momma, T. Y., Kuhnle, G. G., Keen, C. L. & Schroeter, H. Structurally related (–)-epicatechin metabolites in humans: assessment using *de novo* chemically synthesized authentic standards. *Free Radic Biol Med* **52**, 1403–1412 (2012).
31. Taylor, J. I., Grace, P. B. & Bingham, S. A. Optimization of conditions for the enzymatic hydrolysis of phytoestrogen conjugates in urine and plasma. *Anal Biochem* **341**, 220–229 (2005).
32. Xiao, J. F., Zhou, B. & Resson, H. W. Metabolite identification and quantitation in LC-MS/MS-based metabolomics. *Trends Analyt Chem* **32**, 1–14 (2012).
33. Stokvis, E., Rosing, H. & Beijnen, J. H. Stable isotopically labeled internal standards in quantitative bioanalysis using liquid chromatography/mass spectrometry: necessity or not? *Rapid Commun. Mass Spectrom.* **19**, 401–407 (2005).
34. Wang, S., Cyronak, M. & Yang, E. Does a stable isotopically labeled internal standard always correct analyte response? *J Pharm Biomed Anal* **43**, 701–707 (2007).
35. Valleix, A., Carrat, S., Caussignac, C., Léonce, E. & Tchaplal, A. Secondary isotope effects in liquid chromatography behaviour of 2H and 3H labelled solutes and solvents. *Journal of Chromatography A* **1116**, 109–126 (2006).
36. Administration, F. Guidance for Industry - Bioanalytical Method Validation. *Tobacco Smoke Exposure Biomarkers* 247–264, <https://doi.org/10.1201/b18276-17> (2001).
37. Sun, Q. *et al.* Urinary Excretion of Select Dietary Polyphenol Metabolites Is Associated with a Lower Risk of Type 2 Diabetes in Proximate but Not Remote Follow-Up in a Prospective Investigation in 2 Cohorts of US Women. *Journal of Nutrition* **145**, 1280–1288 (2015).
38. Achaintre, D. *et al.* Differential Isotope Labeling of 38 Dietary Polyphenols and Their Quantification in Urine by Liquid Chromatography Electrospray Ionization Tandem Mass Spectrometry. *Anal Chem* **88**, 2637–2644 (2016).
39. Spencer, J. P. E. Metabolism of tea flavonoids in the gastrointestinal tract. *J Nutr* **133**, 3255S–3261S (2003).
40. Rechner, A. R. *et al.* The metabolic fate of dietary polyphenols in humans. *Free Radic Biol Med* **33**, 220–235 (2002).
41. Kuhnle, G. G. Nutrition epidemiology of flavan-3-ols: the known unknowns. *Molecular Aspects of Medicine* (2017).
42. Kuhnle, G. G., Dell'Aquila, C., Runswick, S. A. & Bingham, S. A. Variability of phytoestrogen content in foods from different sources. *Food Chemistry* **113**, 1184–1187 (2009).
43. Hedrick, V. E. *et al.* Dietary biomarkers: advances, limitations and future directions. *Nutr J* **11**, 1–1 (2012).
44. Lampe, J. W. *et al.* Dietary biomarker evaluation in a controlled feeding study in women from the Women's Health Initiative cohort. *Am J Clin Nutr* **105**, 466–475 (2017).
45. Lampe, J. W. & Chang, J.-L. Interindividual differences in phytochemical metabolism and disposition. *Semin. Cancer Biol.* **17**, 347–353 (2007).
46. Setchell, K. & Cole, S. Method of defining equol-producer status and its frequency among vegetarians. *J Nutr* **136**, 2188–2193 (2006).
47. Rodriguez-Mateos, A. *et al.* Influence of age on the absorption, metabolism, and excretion of cocoa flavanols in healthy subjects. *Molecular Nutrition and Food Research* **59**, 1504–1512 (2015).
48. Rowland, I. R., Wiseman, H., Sanders, T. A., Adlercreutz, H. & Bowey, E. A. Interindividual variation in metabolism of soy isoflavones and lignans: influence of habitual diet on equol production by the gut microflora. *Nutr Cancer* **36**, 27–32 (2000).
49. Lampe, J. W. Isoflavonoid and lignan phytoestrogens as dietary biomarkers. *J Nutr* **133**(Suppl 3), 956S–964S (2003).
50. Vought, R. L., London, W. T., Lutwak, L. & Dublin, T. D. Reliability of Estimates of Serum Inorganic Iodine and Daily Fecal and Urinary Iodine Excretion from Single Casual Specimens. *Journal of Clinical Endocrinology and Metabolism* **23**, 1218–1228 (1963).
51. Kuijsten, A., Arts, I. C. W., Vree, T. B. & Hollman, P. C. H. Pharmacokinetics of Enterolignans in Healthy Men and Women Consuming a Single Dose of Secoisolaricresinol Diglucoside. *Journal of Nutrition* **135**, 795–801 (2005).
52. Mennen, L. I. *et al.* Urinary excretion of 13 dietary flavonoids and phenolic acids in free-living healthy subjects - variability and possible use as biomarkers of polyphenol intake. *Eur J Clin Nutr* **62**, 519–525 (2007).
53. Rothwell, J. A. *et al.* Phenol-Explorer 3.0: a major update of the Phenol-Explorer database to incorporate data on the effects of food processing on polyphenol content. *Database* **2013**, bat070–bat070 (2013).
54. Vogiatzoglou, A. *et al.* Flavonoid intake in European adults (18 to 64 years). *PLoS One* **10**, e0128132–22 (2015).
55. Bingham, S. A. *et al.* Validation of dietary assessment methods in the UK arm of EPIC using weighed records, and 24-hour urinary nitrogen and potassium and serum vitamin C and carotenoids as biomarkers. *Int J Epidemiol* **26**(Suppl 1), S137–S151 (1997).

## Acknowledgements

We thank all EPIC-Norfolk study participants and staff for their contribution to the study. We thank all volunteers at UC Davis.

## Author Contributions

J.I.O., H.S. and G.G.C.K. designed the research. J.I.O., R.F., J.K. and J.L.E. conducted the intake escalation and specificity studies. J.I.O., R.F., P.B.G., D.H.M., A.T., A.W. and G.G.C.K. developed and validated the analytical methods. A.B., D.L., P.B.G., D.H.M., A.T., A.W., R.L. and G.G.C.K. analysed and processed samples of the prospective cohort study. K.T.K. is PI of EPIC Norfolk. J.I.O., H.S. and G.G.C.K. had primary responsibility for the final content. All authors read and approved the final manuscript.

## Additional Information

**Competing Interests:** H.S. and J.I.O. are employed by Mars, Inc., a company engaged in flavanol research and flavanol-related commercial activities. G.G.C.K. has received an unrestricted research grant from Mars Inc.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018